

Katherine Goode and Kathleen Rey ††

Introducing ggResidpanel: An R Package for Easy Visualization of Residuals

Introduction and Purpose

As consultants on a wide variety of projects across many majors, a common oversight we encounter is a failure to examine the residuals. We were inspired by the residual panel in SAS to create an R package that easily provides users with a similar panel of plots for 'lm', 'glm', 'lmer', and 'glmer' models using ggplot2. An auxiliary function was also included to receive and plot residuals from any model type. Further capabilities include interactive plots for identifying the data associated with outliers and influential points.

Plots Available

"boxplot": A boxplot of residuals

"cookd": A plot of Cook's D values versus observation numbers

"hist": A histogram of residuals (optional number of bins)

"ls": A location-scale plot of residuals

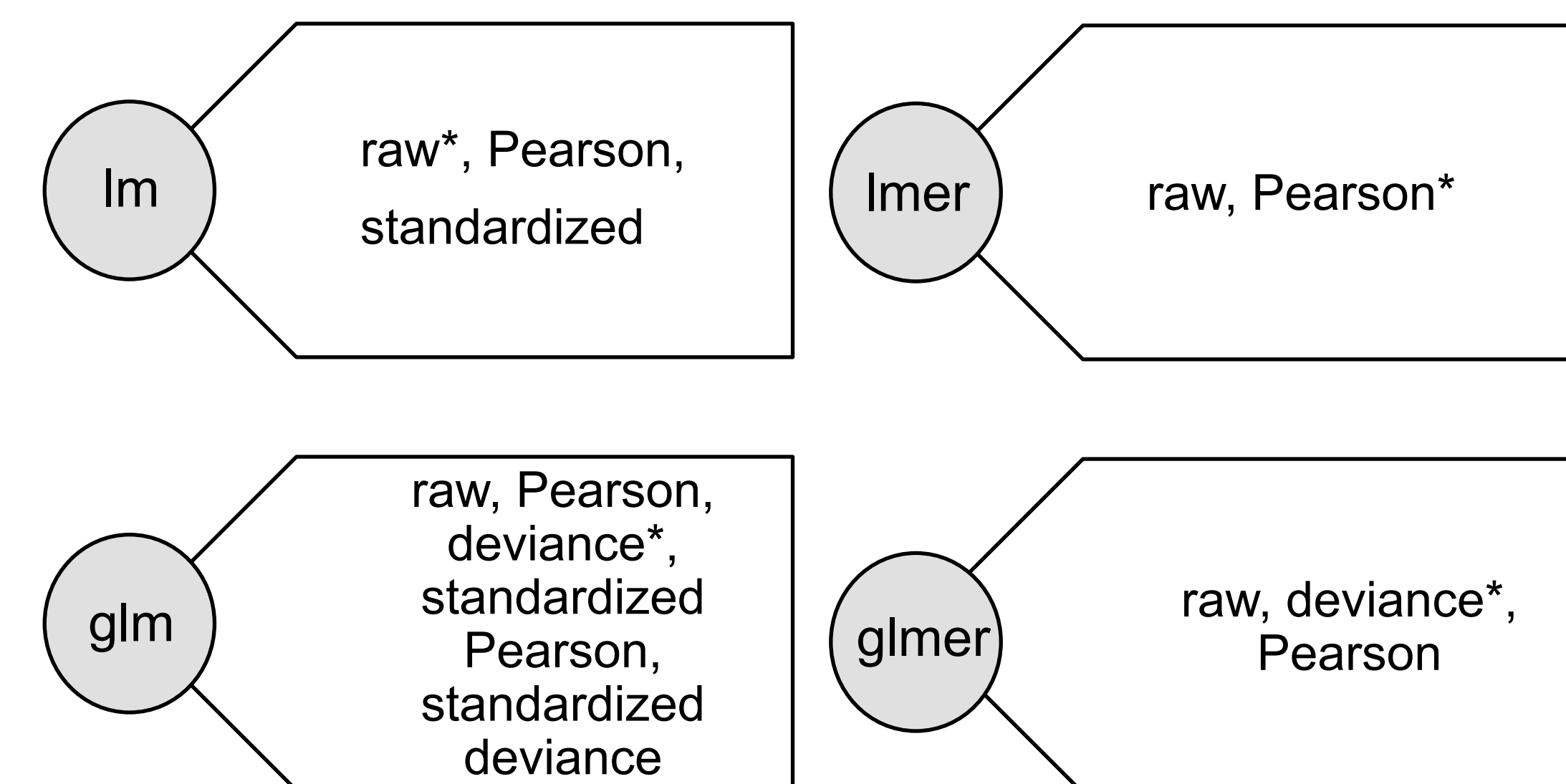
"qq": A normal quantile plot of residuals (optional confidence bands)

"lev": A plot of leverage values versus residuals

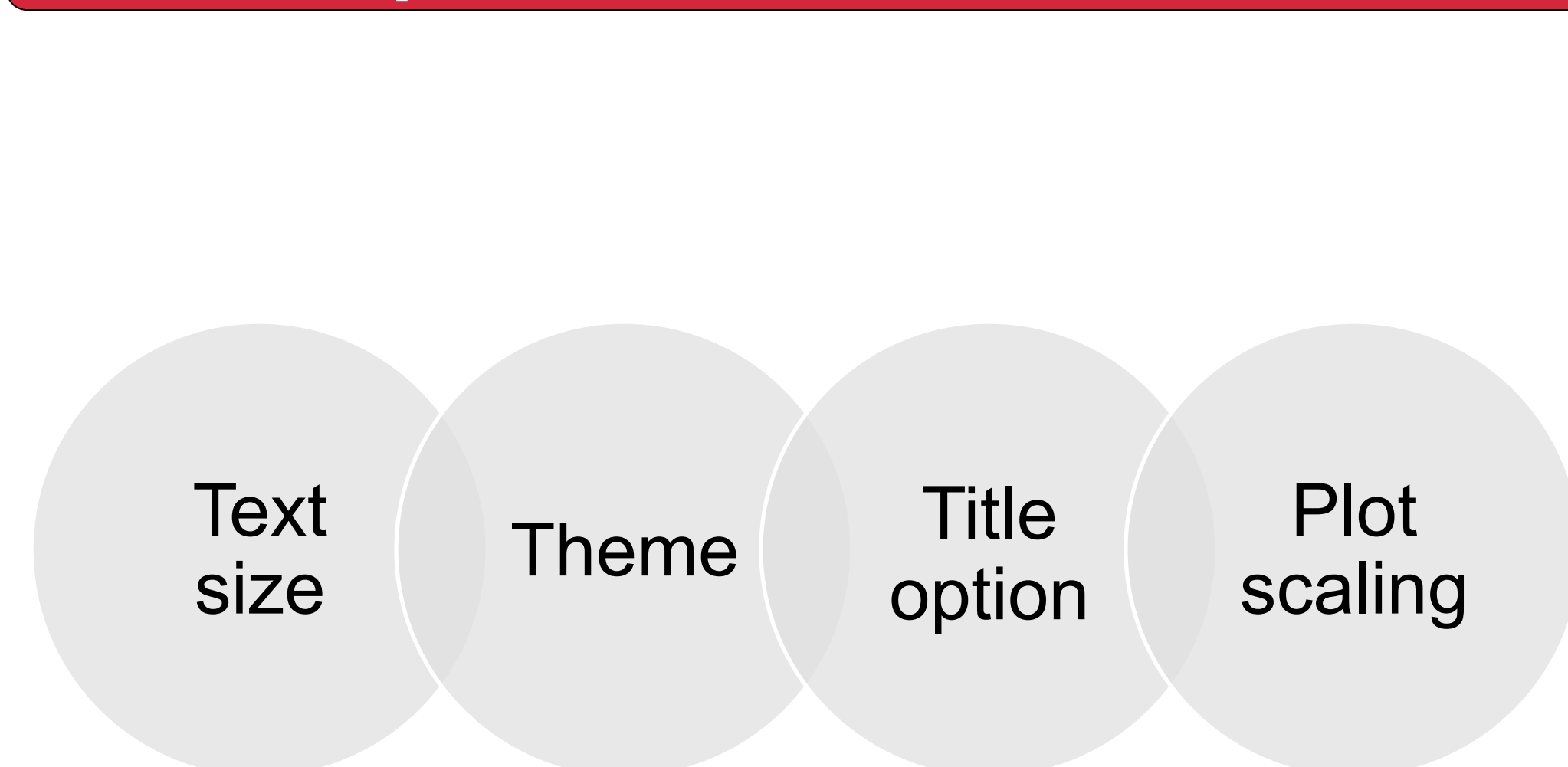
"resid": A plot of residuals versus predicted values (optional smoother)

"yvp": A plot of observed response values versus predicted values

Residual Types (* denotes default)



Format Options



resid_panel: This function is designed to receive a model object and produce a panel of plots to analyze residuals and model assumptions from a variety of viewpoints.

```
resid_panel(model, qqbands = TRUE)
```

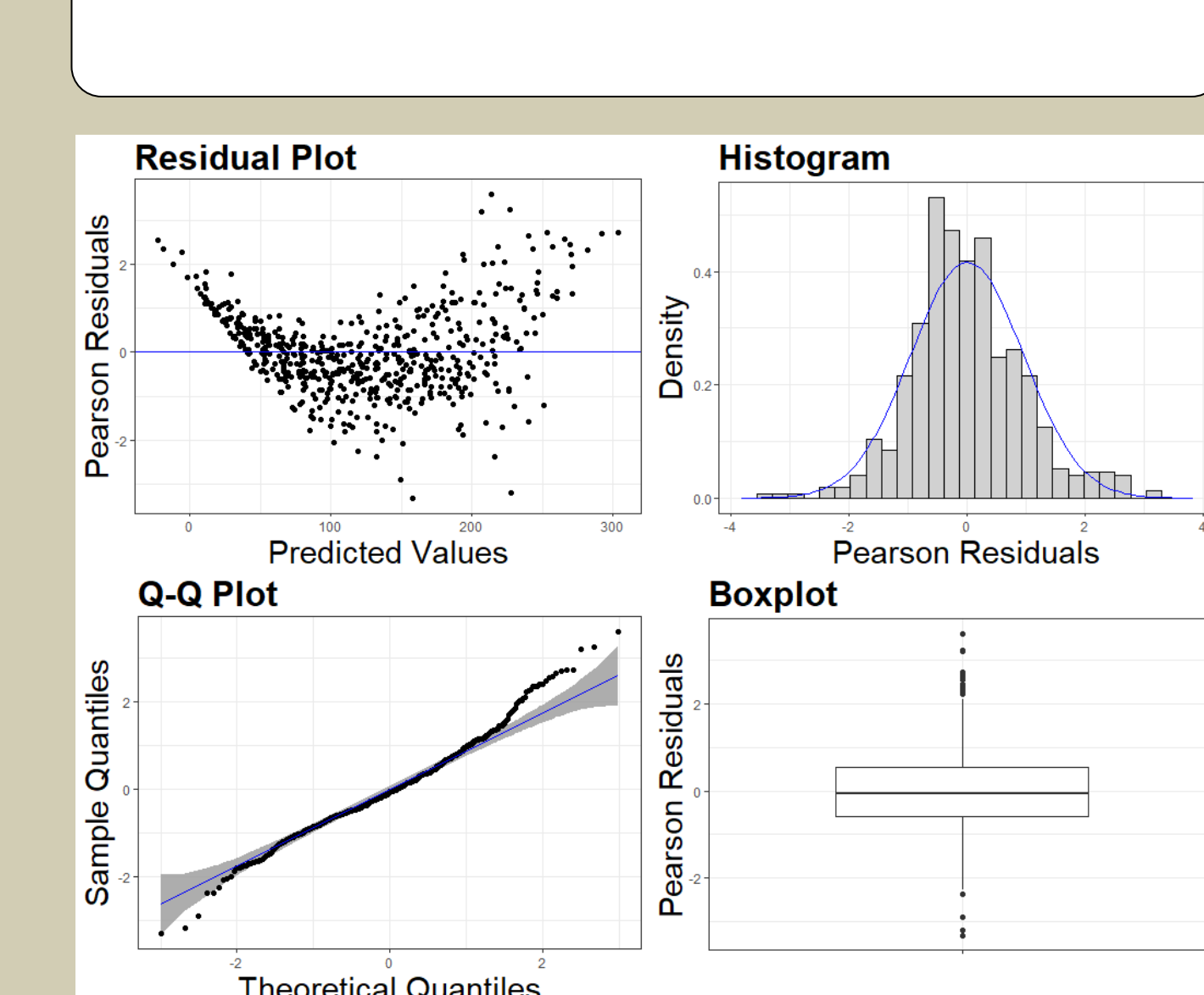


Figure 1: The default panel inspired by SAS

```
resid_panel(model, plots = "R",  
smoother = TRUE, theme = "classic")
```

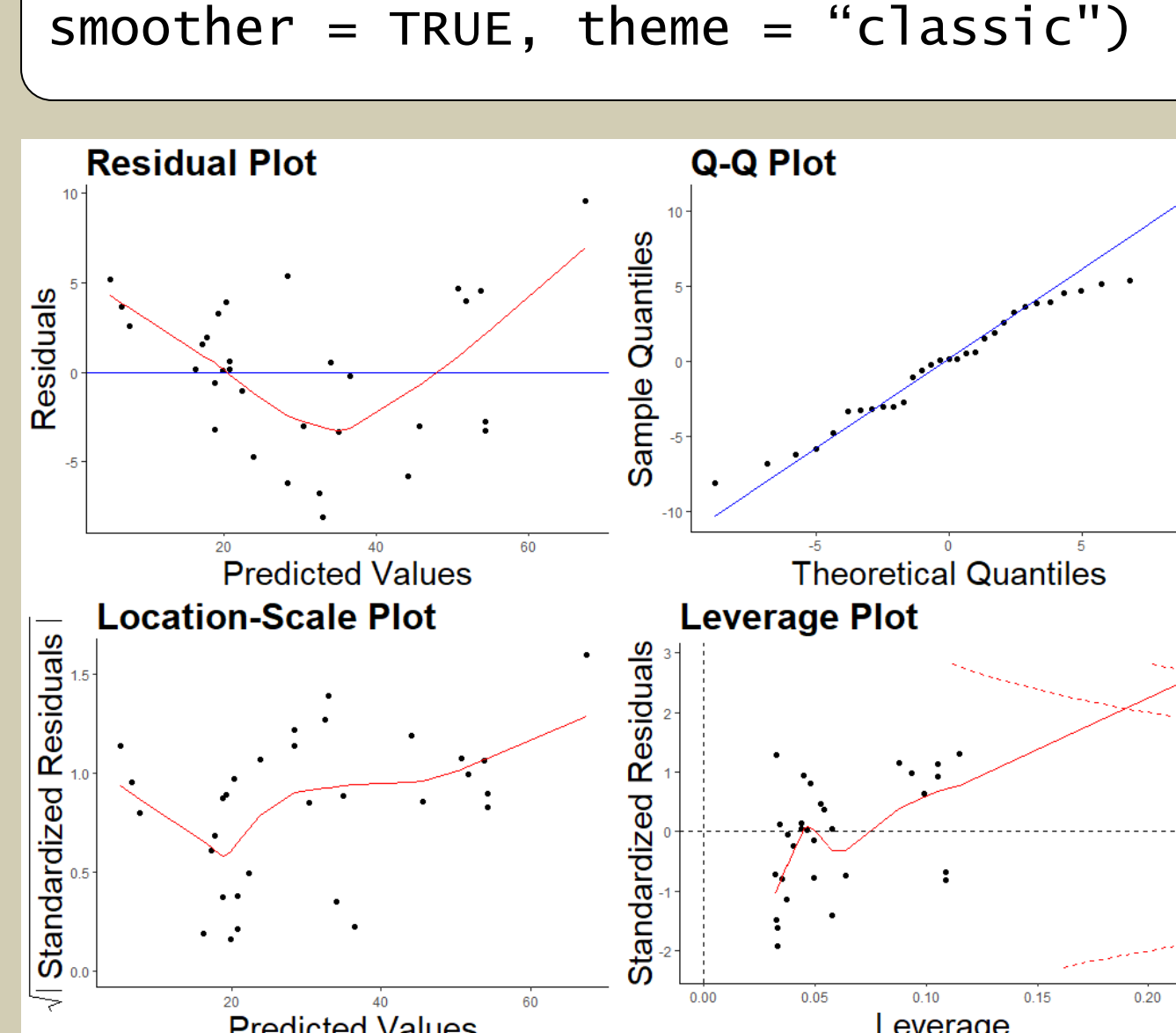


Figure 2: A panel that presents the base R plots

```
resid_panel(model, type =  
"pearson", plots = "resid",  
title.opt = FALSE, smoother =  
TRUE, axis.text.size = 12)
```

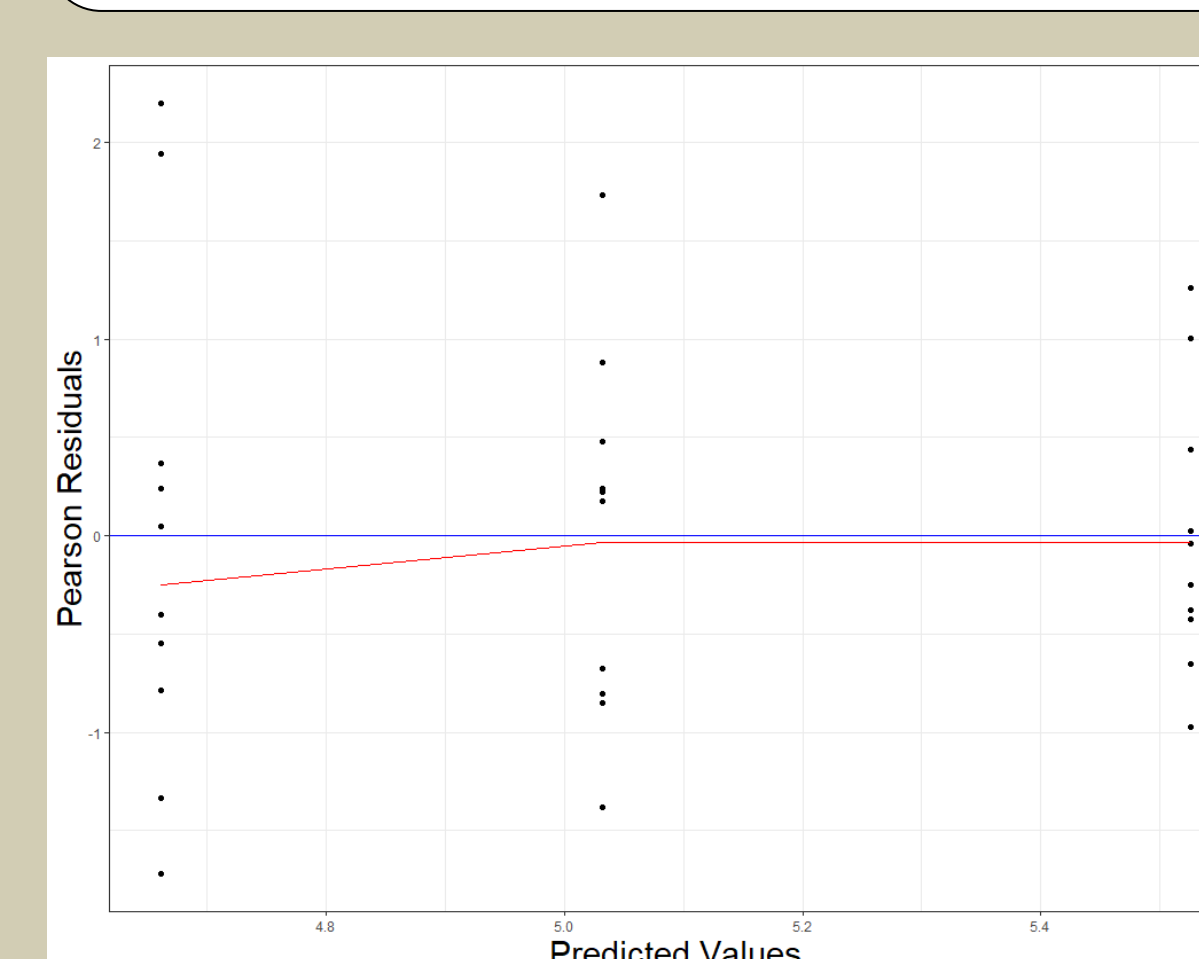


Figure 3: A single plot with a variety of format options

```
resid_panel(model, plots = c("cookd", "qq", "resid"),  
ind.ncol = 3, theme = "grey")
```

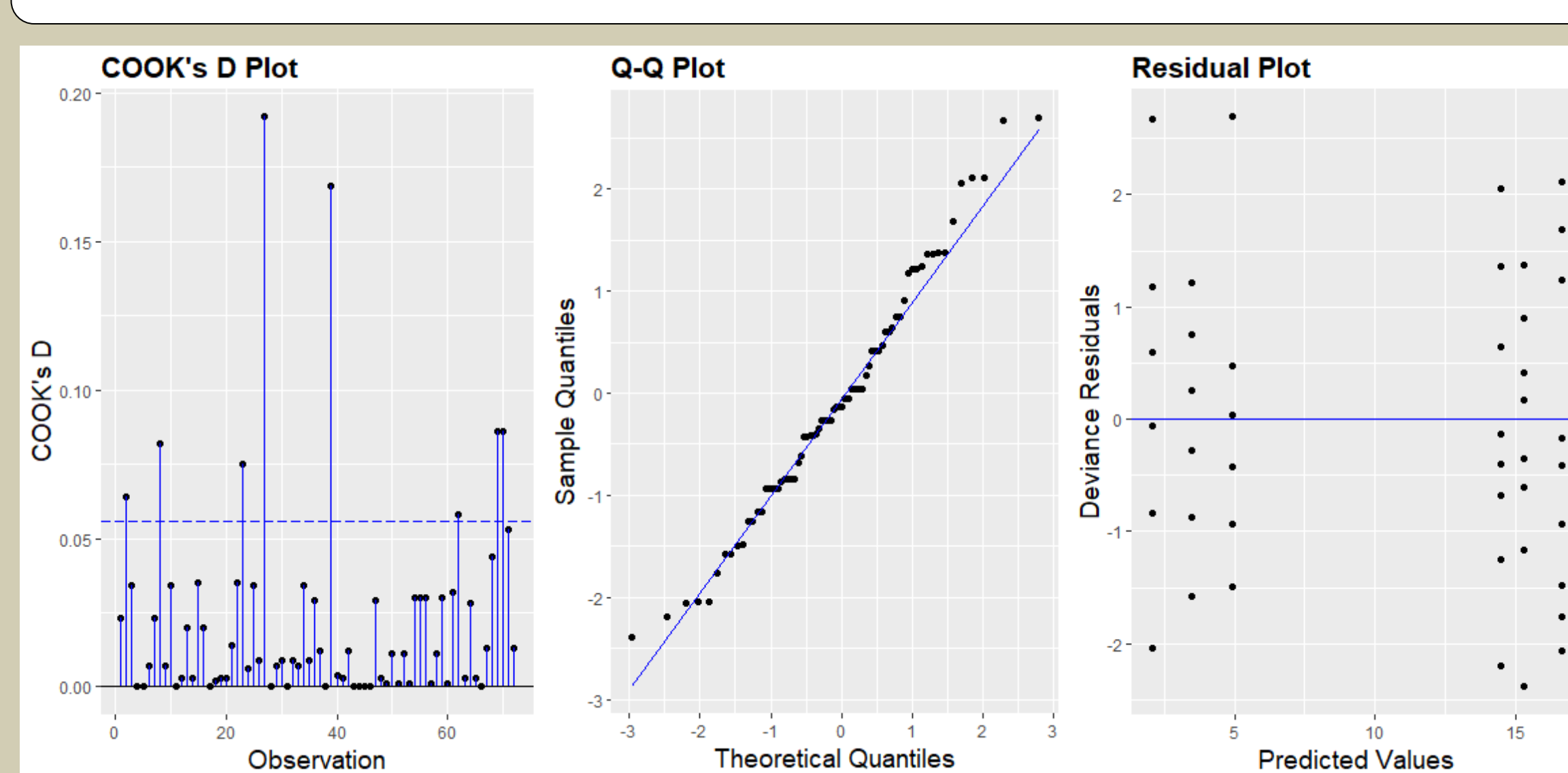


Figure 4: The user can choose which plots to display in a panel

```
resid_panel(model, plots = "all")
```

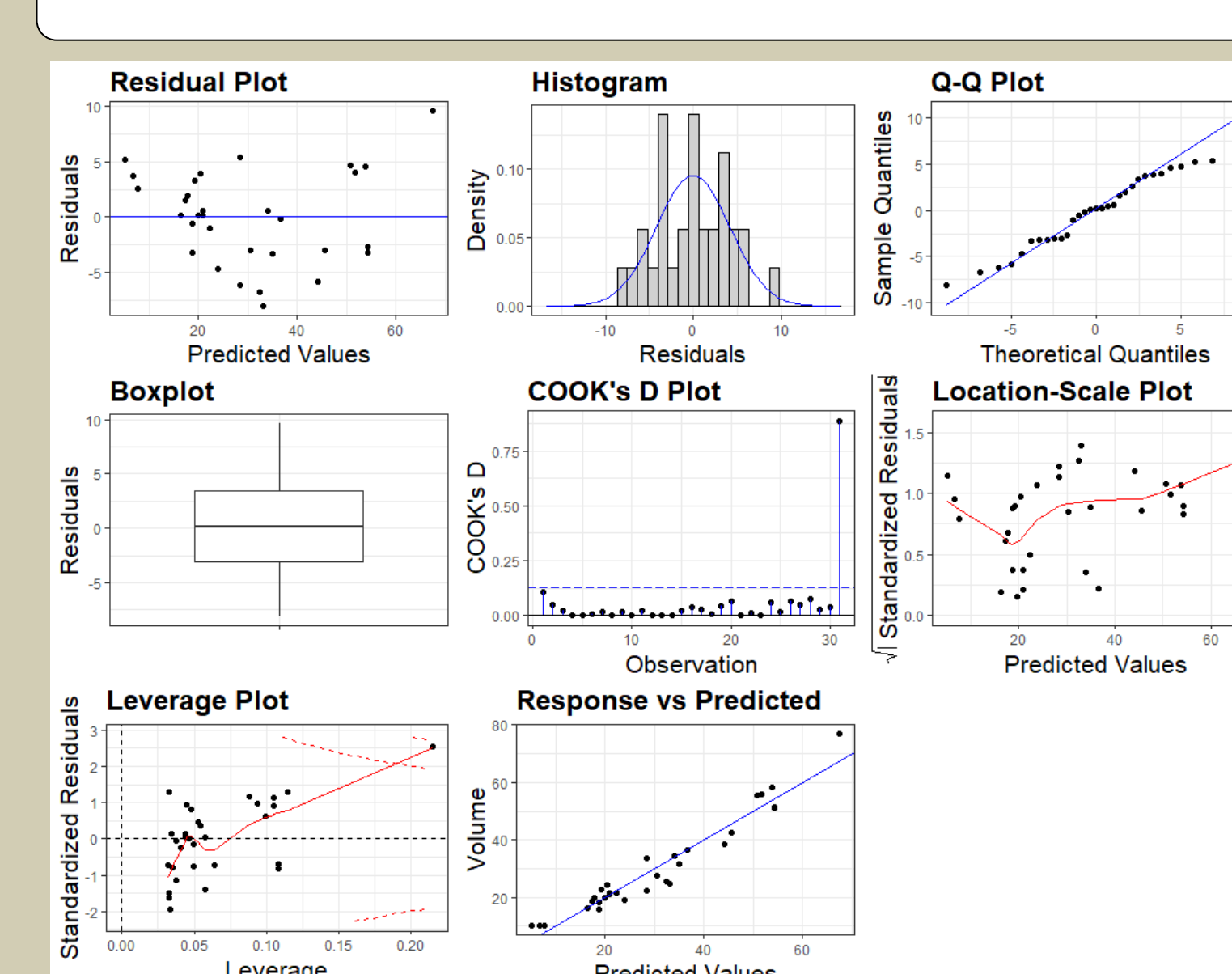


Figure 5: A panel with all available plots

resid_auxpanel: This function is designed to directly receive the residuals which enables the user to assess a wider variety of models.

```
resid_auxpanel(resid(model, type =  
"deviance"), fitted(model))
```

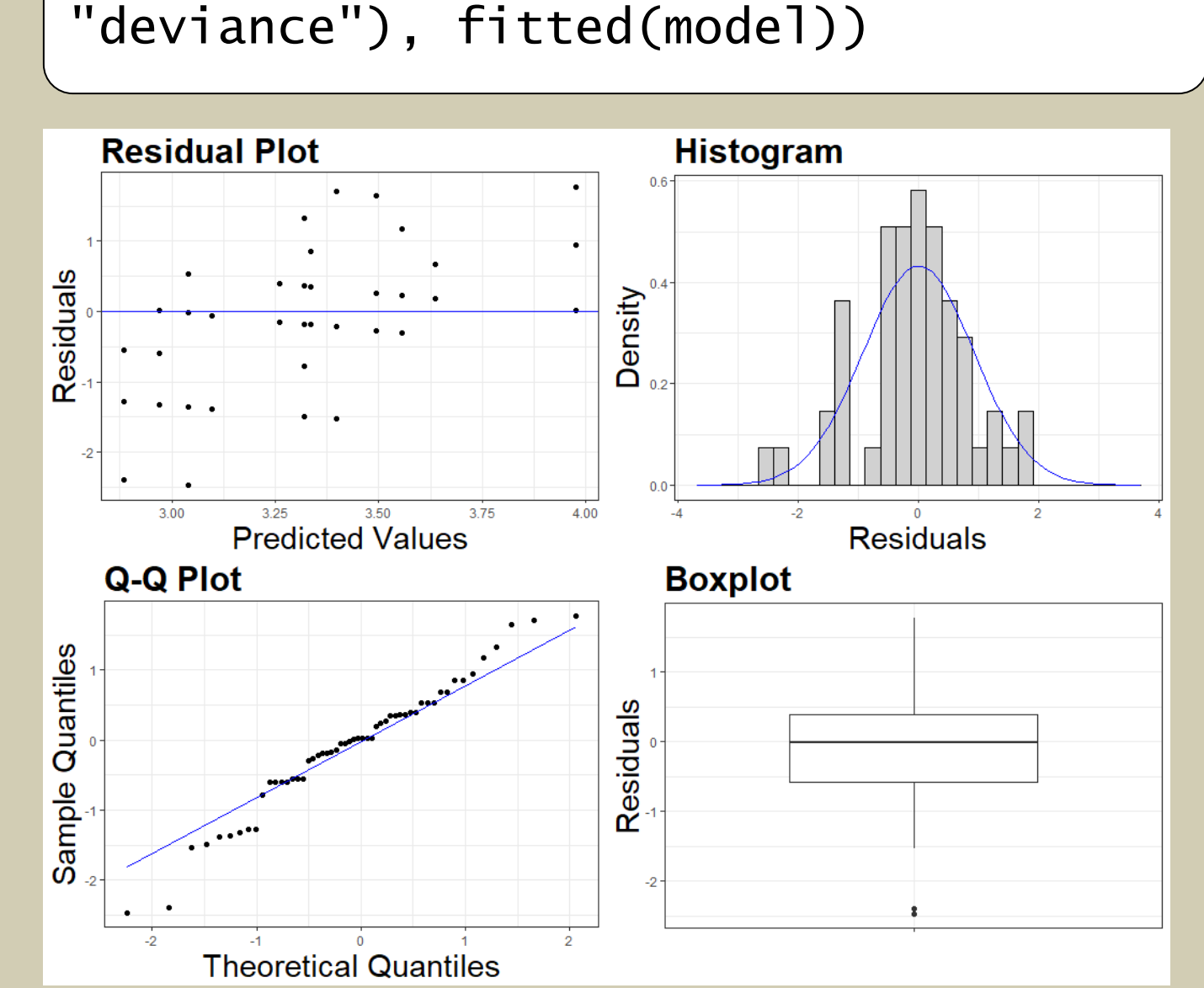


Figure 9: Allows the user to directly input the residuals while producing the same panel as 'resid_panel'

```
resid_auxpanel(resid(model, type =  
"deviance"), fitted(model),  
plots = "qq")
```

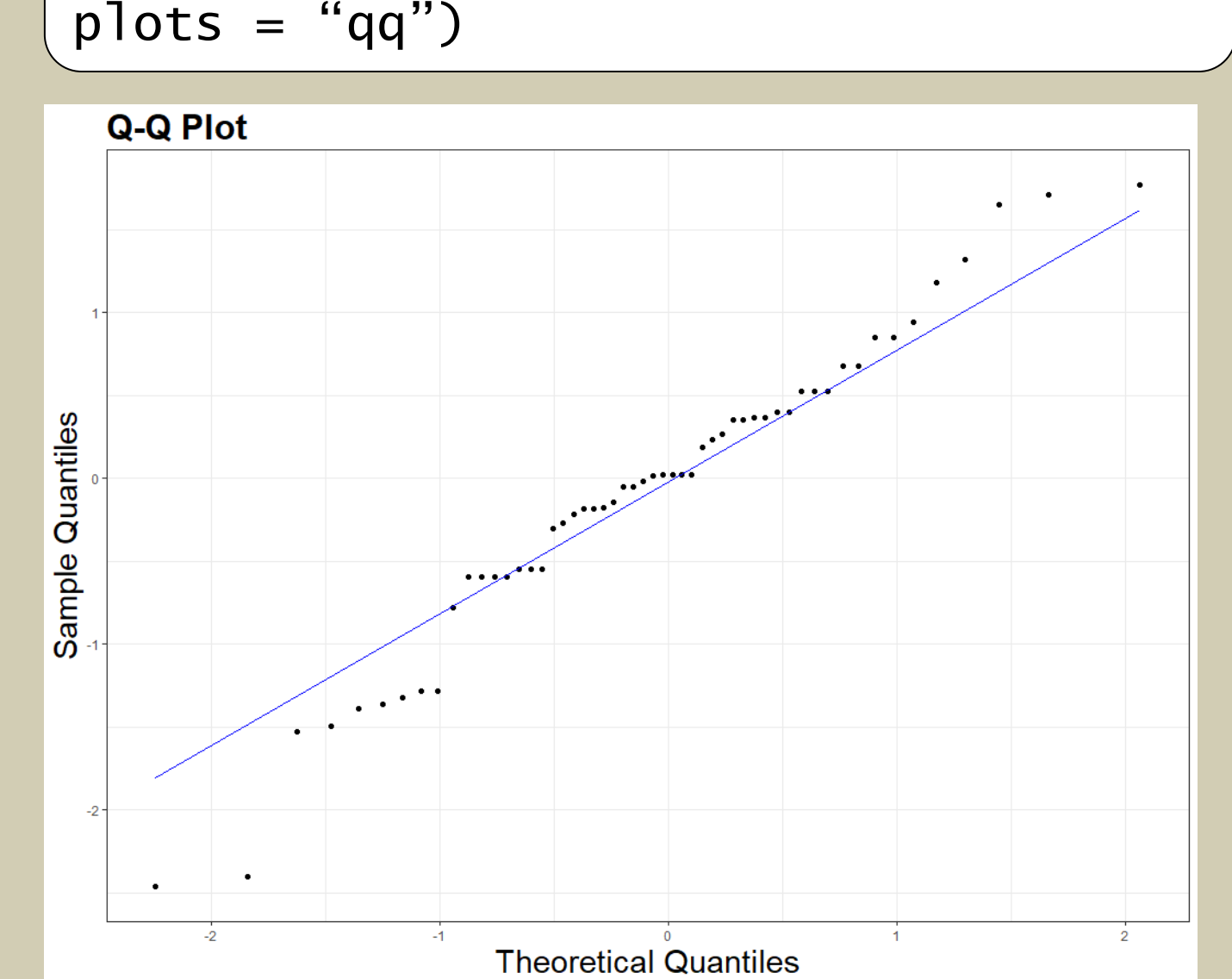


Figure 10: Individual plot selected with 'resid_auxpanel'

resid_interact: This function is designed to connect the data to the plot in an interactive way using 'plotly'.

```
resid_interact(model, plot =  
"cookd")
```

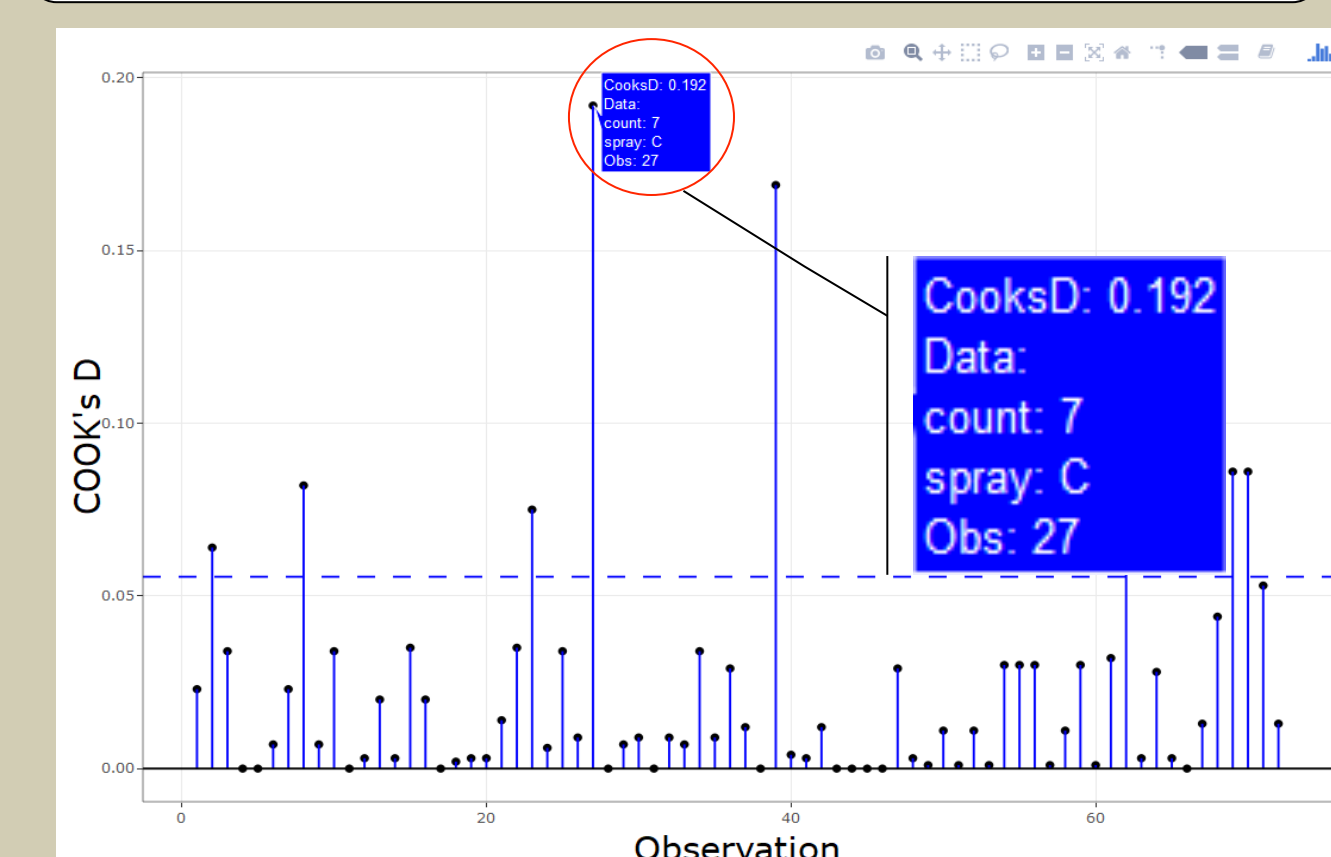


Figure 6: An interactive Cook's D plot

```
resid_interact(model, plot =  
"resid")
```

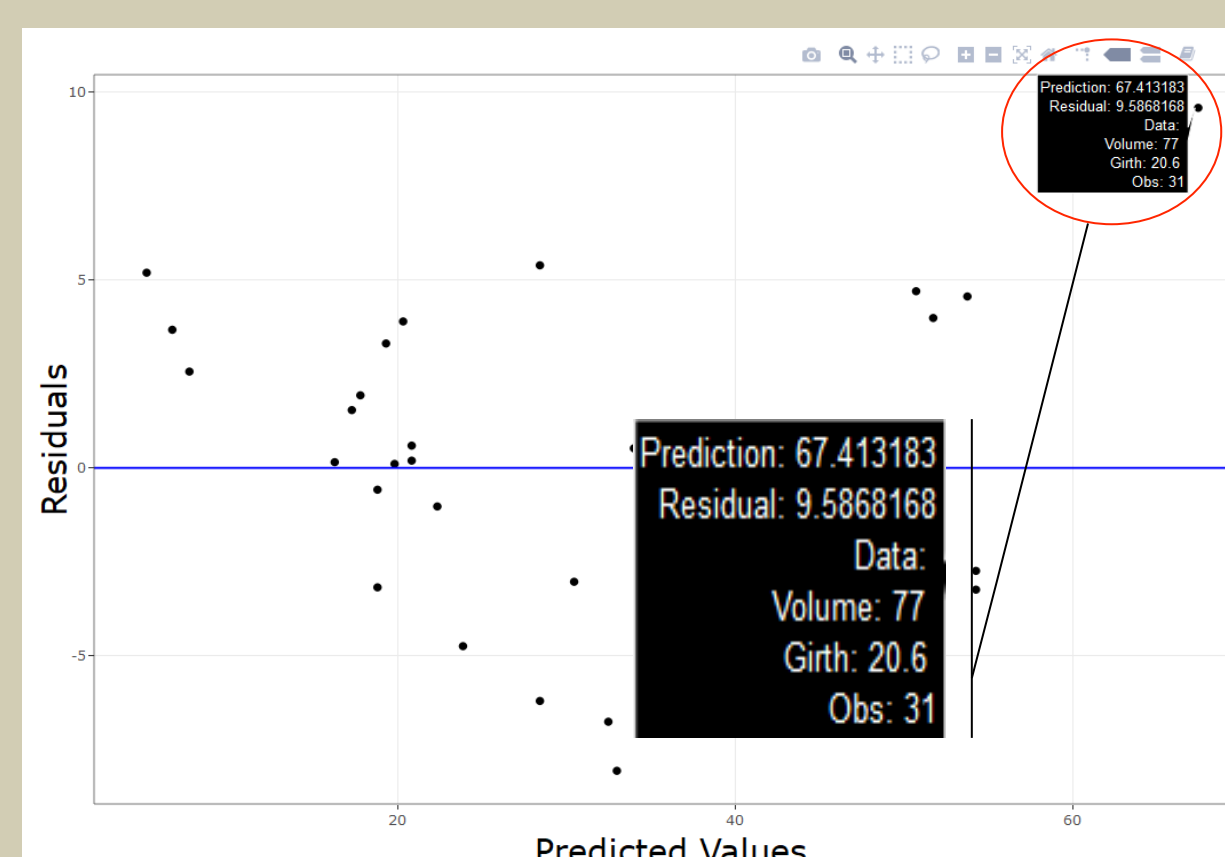


Figure 7: An interactive residual plot

```
resid_interact(model, plot =  
"boxplot")
```

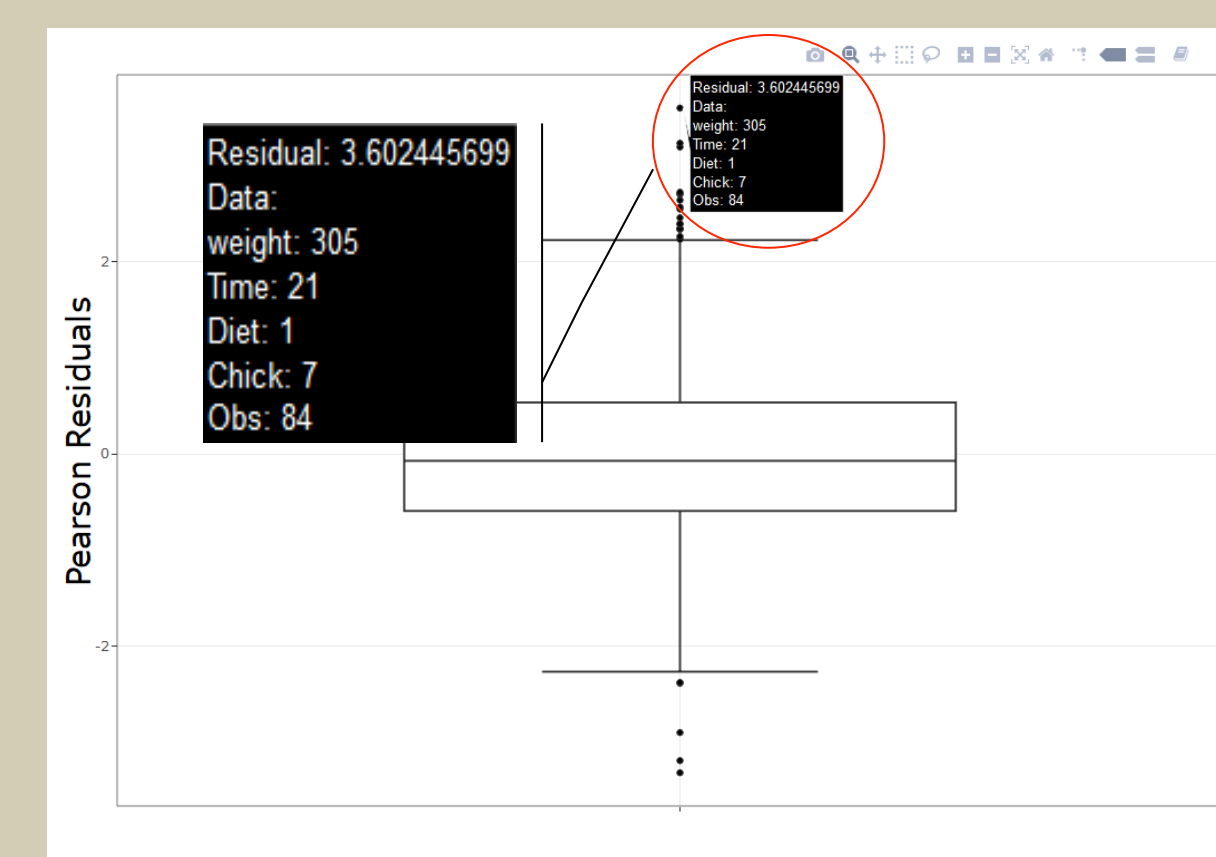


Figure 8: An interactive boxplot

Future Work

- Add a function that accepts multiple models to compare the residuals
- Create a grid of interactive plots
- Incorporate an extended panel similar to SAS that includes model summary statistics
- Integrate an option for a "theme object" to allow users to further adjust the formatting using all capabilities available in ggplot2
- Include an option for standardized residuals with 'lmer' and 'glmer' models
- Write a vignette with more detailed instructions on how to use the package and specifications of how residuals are computed